DOCUMENT RESUME

ED 121 261 IR 003 236

AUTHOR Kiewitt, Eva L.
TITLE Cost-Benefit Analysis of a Computer Retrieval System.
INSTITUTION Indiana Univ., Bloomington. Graduate Library School.
PUB DATE 76
NOTE 22p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage
DESCRIPTORS Computer Programs; *Cost Effectiveness; Costs; Data Bases; *Information Retrieval; Information Services; Information Systems; Program Evaluation; *Relevance (Information Retrieval); Search Strategies; Use Studies
IDENTIFIERS Batch Systems; Controlled Vocabulary; Educational Resources Information Center; ERIC; Indiana University; *PROBE; Text Searching

ABSTRACT

PROBE, a batch-mode information retrieval program used to search the Educational Resources Information Center (ERIC) data base, was developed by Indiana University. To evaluate the effectiveness of the program, questionnaires were given to the users of 200 consecutive searches during February 1973. Relevance ratios were computed by comparing the number of documents considered useful by the user to the number retrieved. For searches using ERIC descriptors, the average relevance ratio was 54.7%. For searches using natural language, it was 42%. It was found that the cost of a descriptor search was $14.42 when 150 searches were run each month and $9.18 per search when 300 searches were run. The cost of abstract searches was $22.25 for 150 searches per month and $17.01 for 300 per month. Based on these findings and a separate user study, it was concluded that the PROBE service was satisfying the user and was cost-effective in that it could become self-supporting at a nominal fee. (Author/PF)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED121261

COST-BENEFIT ANALYSIS

OF A COMPUTER RETRIEVAL SYSTEM

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Eva L. Kiewitt

Graduate Library School

Indiana University

Bloomington, Indiana 47401

2

IR 003 236

COST-BENEFIT ANALYSIS OF A COMPUTER RETRIEVAL SYSTEM

INTRODUCTION

Many libraries and information centers are developing local computer programs to search national data bases. Even though some packaged programs are available, they may be inadequate because of local needs, expense, or computer hardware. Because of spiraling costs and changing needs, continued evaluation must be a component of any program.

One data base that is in heavy demand because of its broad coverage of education and related fields is the tapes from the Educational Resources Information Center. ERIC was established by the U.S. Office of Education (USOE) in 1966 in response to a growing concern for control of proliferating research reports in the field of education. USOE's National Center for Educational Communication (NCEC) managed the ERIC System until 1973, when it was changed from USOE's sponsorship to the National Institute of Education (NIE).

The national ERIC network is composed of specialized subject centers or clearinghouses which select, screen, and abstract educational research reports and disseminate this research by means of microfiche collections and magnetic tapes. The number of clearinghouses has varied from 12 to 20, as some have merged and others have been added. At present there are 16 since the recent merger of Adult Education with Vocational and Technical Education to form Career Education. Also, the

Clearinghouse on Library and Information Sciences joined Educational Media and Technology to become the Clearinghouse on Information Resources.

The two major ERIC index tools, Resources in Education (RIE) and Current Index to Journals in Education (CIJE), are available in hard copy and on machine-readable magnetic tapes. RIE, then Research in Education, was begun in 1966 as a monthly abstract journal with subject, author, and institution indexes to the ERIC research reports. In 1969 CIJE, a journal index, was begun. It currently indexes over 700 journals, more than twice as many as those covered by the Education Index.

In 1968, Burchinal evaluated the first two years of ERIC's development and found significant progress toward four goals:

> to make significant, but previously unavailable, R & D reports easily and readily available to educators.
>
> to interpret and summarize results in ways that educational practitioners and decision-makers can use them.
>
> to help strengthen existing educational communication channels for putting R & D results to use.
>
> to become an important base for developing a national education information network.[1]

A computer information retrieval program, QUERY, was developed in 1969 under the auspices of the U.S. Office of Education for state and local educational agencies and select USOE installations. However, QUERY was found to be uneconomical

because of limited capacity of the IBM System 360/30, with its extreme slowness of operation and its inability to perform complex search specifications. With increased use of ERIC tapes and the unsuitability of QUERY, information centers began to develop their own computer retrieval programs.

THE PROBE PROGRAM

PROBE, a batch-mode retrieval program of the ERIC data base, was developed at Indiana University in 1970 and implemented on the Wrubel Computing Center CDC 6600.* The program has a variable-field format which gives it a broad range of capabilities. A search may be made through any of the following fields: author, title, descriptor (subject heading), identifier (additional non-thesaurus terms), abstract, accession number, sponsoring agency, institution, publication date, dissertation, and others. A combination of these fields may be used, if necessary. Weighted searches may be made by adding a special control parameter to any descriptor or term and by specifying the desired weight and total.

Optional parameters can be set, such as the total number of records to be printed, the number of abstracts for complete printouts, the range of fields to be searched, and the symbol to be used as a string delimiter.

---

*PROBE was developed in 1970 by Ronald Tschudi, who at the time was employed on the ERIC Evaluation Study of the Graduate Library School of Indiana University, Bloomington, Indiana.

Punched cards are prepared for the individual searches, and there is option for a variety of queries. Combinations of Boolean logic search strategies expressed in 'and', 'or', and 'not' combinations are specified. Tschudi and Meredith described the program in more detail in a technical report published in 1972.[2]

During a period of experimentation from January to August, 1971, a free batch-mode retrospective computer search was made available to students and faculty of the Indiana University School of Education. Since the School of Education was receiving a standing order of the ERIC microfiche collection, the users had direct access to the full reports, with microfiche readers and a reader-printer available.

As the demand for searches increased, a $4.00 fee was established in September, 1971, to help defray costs of part-time PROBE personnel. In order to offer the service to all libraries in Indiana, a proposal was presented to the Indiana State Library for the use of grant funds. In 1972 the project was awarded funds provided by the Library Services and Construction Act (LSCA) to publicize and help support the continuation of the computer search for use throughout the state.

A complete description of the experimental phase of the program is explained in an earlier report.[3]

EVALUATION

As the project grew, studies of other retrieval systems

were reviewed to determine the most appropriate evaluation design. Cuadra stated that there are four activities necessary in planning an information system: identifying the information needs of the potential user, predicting the uses to which the information will be put, designing the system that can meet the requirements, and then evaluating and re-evaluating the resulting system.[4] The same basic components are usually considered in both designing and evaluating an information system.

Various methodologies or measures of evaluation have been used to test information retrieval systems. The user must be studied to determine his needs and whether those needs are being satisfied; the system must be tested to see whether performance is efficient and effective; and costs must be ascertained to determine fees and need for future funding. When these studies are completed, the needs and costs must be compared with the objectives to discover whether the goals are being met.

Lancaster stated that we can evaluate an information retrieval or dissemination system from any of the following three viewpoints:

1. How well the system is satisfying its objectives, which will usually mean how well it is satisfying the demands placed upon it. Here we are evaluating the effectiveness of the system.

2. How efficiently (in terms of costs) it is satisfying its objectives. This is cost-effectiveness evaluation.

3. Whether the system justifies its existence (i.e., the system worth). In evaluating system worth, we are concerned with cost-benefit relationships.[5]

Beginning in 1967 the literature concerning the evaluation of information systems has been reviewed in the Annual Review of Information Science and Technology.[6]

Cleverdon stated in his review of 1971:

> Whereas in earlier years most of the tests have been concerned with performance, there has recently been shown an increasing interest in costs and, thereby, in cost effectiveness. . . while there have been no reports of major independent evaluations of operational systems this year, the practice of in-house testing is growing.[7]

Even though there had been an increasing number of evaluation studies, reviewers stated the need for additional careful evaluations. The importance of cost analysis was shown by the addition of a separate chapter on costs and budgeting in the 1972 volume of Annual Review of Information Science and Technology.[8] In 1971 King and Bryant described various methodologies of evaluation, especially relating to cost and benefit factors;[9] and Swanson and Mayer reported on the state-of-the-art of evaluation study in 1975.[10]

Three aspects of the PROBE project were evaluated. The users' expressed needs and satisfaction were studied by means of a critical incident survey to acquire more understanding of the educators served. The user study is described in a recent issue of College and Research Libraries.[11] The performance

of the retrieval system was tested by the ratios of recall and precision to determine its effectiveness, and the costs were analyzed to ascertain its cost-effectiveness.[12]

SYSTEM PERFORMANCE

Various methods have been developed to measure system performance. Although there is no universally accepted evaluation method, one of the most frequently used is a combination of precision and recall ratios, which was chosen to evaluate this system.

Questionnaires were given to the users of 200 consecutive searches during February, 1973. Information for evaluation was derived from the returned questionnaires and from statistical summaries of the computer searches. The data base used was the ERIC Research in Education tapes, which at that time included 59,495 records on file.

The most frequent method of searching the ERIC tapes with the PROBE program had been by using the descriptors listed in the Thesaurus of ERIC Descriptors.[13] Boolean logic was used to combine the selected descriptors (subject headings) to form the computer query, together with the logical connectors 'and' (conjunction), 'or' (disjunction), and 'not' (negation). A second search was conducted through the abstracts using natural language terms or phrases in order to compare the two types of searches.

The basic evaluation of the system performance was ascertained by the use of the precision or relevance ratio which reflects the ability of the system to avoid retrieval

of unwanted items. The user was asked to determine the number of documents in both searches that he considered relevant to his question and these were compared to the computer statistical summary of the total number of documents retrieved. Thus, the precision ratio was determined by the use of the following formula:

$$\frac{\text{the number of relevant documents retrieved}}{\text{the total number of documents retrieved}} \times 100$$

Table 1 illustrates the user's determination of the total number of relevant reports found in the descriptor search, the total number retrieved, and the precision ratio.

One hundred thirty-two (84.6 percent) of the returned questionnaires included data on the number of documents the users considered relevant.

The number of relevant documents ranged from zero to 800, since some considered that no documents were relevant, and others stated that all documents on the printout were relevant. The program parameters allowed printing of a maximum number of 100 citations with abstracts, plus up to 700 additional citations to be printed. The number of retrieved documents also had a wide range of from one document to 800 documents retrieved. Thus, the precision ratio varied from zero to 100 percent.

The average number of relevant documents retrieved per user was 86, and the average of the total number retrieved was 168 documents. Some of the users very carefully checked each document citation and abstract to give the number of relevant

TABLE 1. SUMMARY OF THE PRECISION RATIOS OF DESCRIPTOR SEARCHES

	Number of Relevant Documents	Number of Retrieved Documents	Precision Ratio
Totals (N = 132)	11,392	22,205	54.8
Range	0-800	1-800	0-100
Mean	86.3	168.5	54.78
Median	22.5	91	63.05

documents. Some gave percentages rather than exact numbers, and several considered all of them relevant if they were interested in a broad search for bibliographic citations, current awareness, or information for a dissertation topic.

The overall precision ratio was 54.78 percent. This ratio compares favorably with the 50.4 percent precision ratio found by Lancaster in his evaluation of the national MEDLARS Demand Search Service.[14]

Variation of the precision ratio can be attributed to several factors. Four major factors include the users, the negotiator, the computer program, and the computer.

Table 2, which gives the summary of the results of the precision of the abstract searches, includes the total number of relevant reports, those retrieved, and the precision ratio.

For the abstract searches the approach was by natural language terms or phrases. One hundred five respondents (94.6 percent) of the 111 returning questionnaires completed the question requesting the number of relevant reports found in the abstract search.

The range of relevant documents found varied from zero to the total maximum allowed of 800 documents. The number of retrieved documents ranged from two to the maximum of 800. Thus, the precision ratio also varied from zero to 100 percent. The average number of relevant documents per user was 86, with an average of 210 documents for each user.

TABLE 2. SUMMARY OF THE PRECISION RATIOS OF ABSTRACT SEARCHES

	Number of Relevant Documents	Number of Retrieved Documents	Precision Ratio
Totals (N = 105)	9,046	32,594	42.0
Range	0-800	2-800	0-100
Mean	86.6	210.4	42.0
Median	27.0	184.0	27.0

As was expected, the overall precision ratio of 42 percent was lower than the descriptor precision ratio of 54.8 percent. With the use of natural language, usually additional minor and nonrelevant documents are obtained, since the search is broader and the phrases used may have a variety of meanings according to the context in which they are used in the abstract.

Recall or sensitivity is a measure of evaluation of the ability of the system to retrieve all relevant items in the file. The number of relevant documents retrieved and the number of relevant documents that ought to be retrieved are combined to form a recall ratio.

It is quite difficult to determine what is relevant, however. Usually, the user is asked to make a value judgement as to whether it is relevant to his request. The size of the system may pose another problem. If there is a large number of documents in the system, it is difficult to assure that all relevant documents are retrieved. In a small set or simulation model, all documents may be examined individually, but this is not possible when studying a large system; therefore, recall is not accurately measureable.

It was decided to use an experimental method to test recall by comparing the number of relevant documents retrieved by a descriptor search to the number of relevant documents obtained by an abstract search. The abstract search was expected to be broader and retrieve additional minor relevant documents. The

following formula used for this simulated recall measure was:

$$\frac{\text{the number of relevant documents retrieved by using ERIC Thesaurus descriptors}}{\text{the number of relevant documents retrieved by natural language terms in the abstracts}} \times 100$$

The total number of computer hits for each of the 200 descriptor searches were compared to the abstract searches for the same user. One hundred fifty-five (77.5 percent) of the paired searches showed more hits from abstract search; forty-three (21.5 percent) of the searches showed more hits from the descriptor search; and 2 (1 percent) had the same number of hits in both searches. The overall average recall ratio of the useable paired searches was 44.4 percent. Thus, this method of computing the recall ratio did not prove entirely satisfactory, and from these results, one cannot assume that the abstract search will always locate additional relevant documents.

COST STUDY OF THE PROBE SERVICE

The cost of PROBE as an information service was studied to determine the pricing needed to obtain financial self-sufficiency. This could be determined by comparing operating expenses with anticipated income, as is shown in the explanation of the cost formula. However, only the actual costs were included; the university absorbs some costs as a research service to students and faculty.

Three types of costs were evaluated: fixed costs and

two types of variable costs. The fixed costs included salaries, computer tapes, and other recurring expenses. One type of variable was the cost relating to the number of searches conducted, i.e., the larger the volume of searches, the less the cost of the individual search. Another type of variable cost was determined by the type of search conducted: the abstract search is usually much more expensive than a descriptor search since it takes more computer time.

In retrospective searches the cost is determined by the number of documents in the data base, the number of searches completed, and the average computer time per search.

A comparison of the data of the descriptor and abstract field searches is shown in Table 3.

After computing the time and costs of the computer searches, the other operating costs were compiled. Costs were divided into three categories: personnel, equipment and supplies, and computer costs.

King's formula, which included both fixed and variable costs, seemed most feasible:[15]

$$C = C' + C'' X_1 + C''' X_2$$

Thus, the equation was cost = fixed costs + variable costs relating to the number of searches conducted + variable costs relating to the mode of searching (descriptor or abstract). Since costs decrease as the number of searches increases, two cost estimates were given: one for 150 searches per month and another for 300 searches per month. Computer costs were determined by a formula

TABLE 3. COMPARISON OF DESCRIPTOR AND ABSTRACT FIELD SEARCHES

Search Block Record Per Average Search	Descriptor Field Searches	Abstract Field Searches
Number of queries	2.4	1.4
Number of specifications (keywords)	8.4	5.3
Number of records printed (abstracts plus citations)	94	395
Number of CP seconds required	40.23	143.22
Cost at commercial rate	$ 6.44	$ 22.21
Cost at university rate	$ 3.15	$ 10.98

used by the Wrubel Computing Center (WCC), in which they charge $2.77 per megaword second. The figures obtained from the computer statistical summaries and information obtained from the WCC were based upon 200 searches processed in three batches.

It was found that the cost of a descriptor search is $14.42 when 150 searches are run each month, and $9.18 per search if there are 300 searches conducted each month. The cost of abstract searches is $22.25 for 150 searches per month, and $17.01 per search if 300 searches are run each month.

The variable computer costs may be decreased using alternate forms of output, such as limiting the number of abstracts printed, searching only the most recent tapes, or batching a larger number of searches for each computer run.

The users were requested to indicate on a questionnaire their willingness to pay. One hundred twenty-eight users recorded the range of prices that they were willing to pay for the PROBE search. Eighty-seven (68 percent) of the users were willing to pay from $5 to $10 per search, and 41 users (32 percent) stated their willingness to pay more than $10 per search. Thus, the majority were quite willing to pay within the present price range, and approximately one-third were willing to pay from $10 to $60.

CONCLUSION

The user was proven to be satisfied with the results of his search since 79.5 percent of the respondents indicated that

their search was of value to them, as shown in the article describing the user study.[16] The three greatest values to the user were the time saved, the speed of the receipt of the search, and comprehensiveness of the searches.

In evaluating the system performance, the precision test was satisfactory overall at a 54.7 percent ratio for descriptor searches and a 42 percent overall ratio for abstract searches. The overall average recall ratio of useable paired searches was 44.4 percent. However, the range was broad and more extreme than desirable. This seems to indicate that the negotiator and user must interact more closely to determine the preferred results of the search, and the user must be made more aware of the capabilities and limitations of the system.

The variation in precision and recall ratios did not alter the fact that the study of the users and system performance showed that the PROBE computer search of the ERIC tapes was reaching its objectives by effectively satisfying the user. It was found that it was possible to continue the service within the cost range that users were willing to pay. Also, the cost factor could be controlled, and if necessary, decreased by limiting the number of abstracts or citations printed, by charging for additional output per abstract, or by searching only the most current tapes.

Based on the findings of the study, it was concluded that the PROBE service was satisfying the user and was cost-

effective in that it could become self-supporting at a nominal fee. As a result of the study, the negiotator structured the queries more precisely to obtain the most relevant documents. Also, the fee was increased from $4.00 to $6.00 per search for Indiana University patrons and out-of-state users were charged $12 in order for PROBE to become self-supporting. Further details of the evaluation may be obtained from the author's complete description of the study.[17]

REFERENCES

1. Lee G. Burchinal, Evaluation of ERIC (Washington, D. C., U. S. Office of Education, ED 020 449, 1968), p. 4.
2. Ronald R. Tschudi and Joseph C. Meredith, The "PROBE" Retrieval Program - A Description (Bloomington, Indiana, Indiana University, ED 059 596, 1972).
3. Eva L. Kiewitt, PROBE Computer Search of the ERICtapes (Bloomington, Indiana, Indiana University, ED 070 050, 1973).
4. Carlos A. Cuadra, "Planning Information Systems and Services," in Annual Review of Information Science and Technology, vol. 6 (Chicago: Encyclopaedia Britannica, 1971), p. 1.
5. F. Wilfrid Lancaster, "The Cost-Effectiveness Analysis of Information Retrieval and Dissemination Systems," Journal of the American Society for Information Science 22:12 (January-February, 1971).
6. Annual Review of Information Science and Technology (Washington, D. C.: American Society for Information Science, 1966-).
7. Cyril W. Cleverdon, "Design and Evaluation of Information Systems," in Annual Review of Information Science and Technology 6:60 (Chicago: Encyclopaedia Britannica, 1971).
8. John H. Wilson, Jr., "Costs, Budgeting, and Economics of Information Processing," in Annual Review of Information Science and Technology 7:39-67 (Washington, D. C., American Society for Information Science, 1972).

9. Donald W. King and Edward C. Bryant, The Evaluation of Information Services and Products (Washington, D. C.: Information Resources Press, 1971).

10. Rowena W. Swanson and Joseph Mayer, "Performing Evaluation Studies in Information Science," Journal of the American Society for Information Science 26:3 (May-June, 1975), pp. 140-156.

11. Eva L. Kiewitt, "A User Study of a Computer Retrieval System," College and Research Libraries 36:7 (November, 1975). (to be published)

12. Eva L. Kiewitt,"PROBE: Computer Searches of the ERICTAPES--An Evaluation of a Pilot Study", unpublished Ph.D. dissertation (Bloomington,Indiana: Indiana University, 1973).

13. Thesaurus of ERIC Descriptors (New York: CCM Information Corporation, 1972).

14. F. Wilfrid Lancaster, Evaluation of the MEDLARS Demand Search Service (Washington, D. C.: National Library of Medicine, ED 022 494, 1968), p. 39.

15. Donald W. King and Edward C. Bryant, The Evaluation of Information Services and Products, op. cit., p. 74.

16. Eva L. Kiewitt, "A User Study of a Computer Retrieval System," op. cit.

17. Eva L. Kiewitt,"PROBE: Computer Searches of the ERICTAPES--An Evaluation of a Pilot Study", op. cit.